1.选题意义与应用背景

随着互联网的普及和Wifi技术的发展，无线/移动通信网络已经成为当今社会的基础设施，广泛地应用于政府、金融、军事等领域。统计显示，截止到2017年12月，中国网民规模达7.72亿，互联网普及率达到55.8%【2018年第41次中国互联网络发展状况统计】；与此同时，随着支付宝和微信支付等线下手机支付方式的普及，我国网民在线下消费中使用手机支付比例由2016年的50.3%提升至65.5%。无线网络设备的普及在给人们生活带来便利的同时，也存在很多安全隐患，如隐私泄露、病毒入侵等，因此对网络安全的防御和无线网络信息安全的需求日益迫切。

保障无线网络安全的传统技术手段是使用密钥进行身份认证，即用户采用账号和密码的方式登陆系统或进行支付等操作。这种方法的缺陷在于，一旦密钥发生泄露将会造成安全漏洞。值得庆幸的是，攻击者可以通过多种方式获取用户的账号密码，甚至模拟用户的操作，但却很难伪造出与用户相似的设备特征。基于此，研究人员提出可以在密钥认证的基础上，根据用户设备的特征再次认证，从而准确识别用户。这样即使用户的账号被盗，攻击者仍然无法登陆用户的账号，极大地增强了用户的信息安全。

2.相关工作

“指纹”最早是指标识个体的生物特征识别技术【**Accelprint\_NDSS14**，Information fusion in biometrics】【P. Tuyls and J. Goseling,“Capacity and examples of templateprotectingbiometric authentication systems,”】。这一概念早在上世纪60年代就被应用于设备识别，当时研究者们开发出可观测到信号特征的系统来区分雷达【K. Talbot, P. Duley, and M. Hyatt, “Specic emitter identification andverification,”】。此后类似的技术被用于蜂窝网络中的发射器识别【M. Reizenman, “Cellular security: better, but foes still lurk,”】【L. Langley, “Specific emitter identification (sei) and classical parameter fusion technology,”】。近年来，随着对无线网络设备研究的深入，对无线/移动网络设备“指纹”的分析技术也越来越受到研究者的重视，展现了广阔的应用前景【】。

设备指纹是指可用于唯一标识出该设备的设备特征或者独特的设备标识。目前较普遍的技术是针对设备提供的某些信息，从中提取特征生成近似唯一的设备指纹，与存储的可信信息匹配辨识，从而对设备的身份进行认证和辨识。其原理是不同厂商生产的设备（即便是同一厂商生产的设备）会由于硬件或软件的不一致在某些特定的方面存在些许差异。当前研究的技术手段可大致分为基于硬件的识别和基于软件的指纹识别。基于软件的识别通常是从流量信息中获取特征指纹，而基于硬件的识别则往往是根据设备的传感器信息或时钟偏移等硬件设备的信息对设备进行标识。

2.1基于软件的设备识别

802.11 MAC帧的格式和内容经常被用于辨识无线设备。MAC地址是Medium/Media Access Control地址的简称，表示互联网上每一个站点的标识符。在网络底层的物理传输过程中，是通过物理地址来识别主机的，所以MAC地址就如同居民身份证上的身份证号码一样，具有全球唯一性。Guo等人根据MAC帧头的信息追踪设备【Do You Hear What I Hear? Fingerprinting Smart Devices Through Embedded Acoustic Components，参考文献40】。顾杨等人利用从MAC层中的管理帧提取的特征指纹作为区分真假AP的特征指纹【】。他们将特征指纹又分为基于无线端的特征指纹和基于特种MAC帧刺激响应的特征指纹，其中基于无线端的特征指纹主要从信标帧和探测响应帧中提取，而基于特种MAC帧刺激相应的特征指纹主要从探测响应帧、关联响应帧和认证帧中提取。第一种特征指纹是通过被动的扫描无线网络获得的，第二种特征指纹是利用构造特殊格式MAC帧主动刺激AP（或无线路由器），捕获AP（或无线路由器）响应获得的。实验结果表明这两类无线设备的特征指纹可以检测出目前大部分无线钓鱼AP。但是，MAC地址是与网卡对应的，如果一个设备拥有不止一个网卡，这个设备就会有相应数目的MAC地址，此时MAC就不能用来唯一地标识设备；另外，如同居民的身份证容易被人盗用或者伪装，设备的MAC地址也容易被篡改或者伪装。

鉴于使用MAC的识别技术中存在上述缺陷，研究者们提出了一系列根据其他特征构建指纹的方法。Desmond等人仅通过研究802.11的请求探测帧，即可识别出连接在同一无线接入点（WLAN:）上的设备。【Do You Hear What I Hear? Fingerprinting Smart Devices Through Embedded Acoustic Components ，参考文献30】。Loh等人通过分析无线局域网（WLAN：Wireless Local Area Networks）中不同设备对802.11请求探测帧响应的时序特性，从中提取相应的指纹【Identifying Unique Devices through Wireless Fingerprinting】。Pang等人也能从流量中提取出特征指纹，进行设备的认证与标识【Do You Hear What I Hear，参考文献57】。Seika【Active Fingerprinting of 802.11 Devices by Timing Analysis】等人向802.11设备发送一个信号，捕获响应帧的到达时间并用时序分析研究其规律，用SVM分类器建立目标设备模型。~~他们在5台设备上测试该方法并达到86%的平均精度~~。

Gao K等人【A passive approach to wireless device fingerprinting】提出了一种用于确定接入到网络的AP的类型的基于黑盒的被动式识别技术。作者进行了大量的实验（收集超过60GB的数据）以对6种AP进行分类，在至少100000个数据包的数据基础上取得了较高的分类精度。无线网络中对AP的识别通常利用网络协议中常见的标识符，例如网络名称、MAC地址或者IP地址，但是这些标识符很容易被伪造、拦截或者更改。

此外，一些开源的工具如Nmap【Do You Hear What I Hear，50】和Xprobe【Do You Hear What I Hear，68】可以通过TCP/IP协议栈的响应识别设备的操作系统。其他基于软件的指纹识别技术则是根据设备上的应用如浏览器来确定设备。Yen等人【同上69】曾经分析长达一个月的必应和Hotmail的日志，成功追踪到用户。Acar等人通过研究JavaScript和一些较流行的第三方插件如Flash Player来获取设备上的字体列表，以此追踪用户【同上，18】。其他研究人员提出使用性能基准来区分JavaScript引擎【同上，54】，浏览器的浏览历史也被用于追踪web端的用户【同上，56】。

2.2基于硬件的设备识别

Bratus等人【Active behavioral fingerprinting of wireless devices】提出了一种主动式设备指纹识别方法,该方法通过向802.11无线设备发送一系列的经过特殊构造的某种错误格式的数据帧来观察设备的回应，通过这些回应的差异可以区分设备在芯片、固件或者驱动上的差异，以此达到设备识别的目的。同时，该方法也可以借助无线网卡芯片或无线网卡驱动的特征指纹信息来区分无线钓鱼AP和合法授权AP，为以较低成本来部署主动式无线端的检测方法提供了一些思路。

Radhakrishnan等人设计了一种名为GTID的识别框架，该技术采用主动和被动结合的方式，不仅可识别物理设备，还可以判断出相应的设备类型【GTID：A Technique for Physical Device and Device Type Fingerprinting】。GTID实现的主要依据是设备之间的异构性，他们认为不同的设备的处理器、DMA控制器和时钟偏移等都存在差异，而这些差异可以被用于识别设备和设备类型。GTID通过被动地抓取网络流量，从中提取出相应的指纹特征并使用人工神经网络（ANN）的算法进行训练和测试，取得了良好的效果。他们的技术适用于各种网络协议，且不需要进行深度包检测，但是由于该技术依赖于细粒度包时间，在路由器和交换机的缓冲区有时间丢失，故该技术的还有待进一步的研究。

【Passive Data Link Layer 802.11Wireless Device Driver Fingerprinting参考文献21】中，作者根据模拟信号特征认证设备。这种指纹识别技术是基于设备的制造和硬件组件在数字领域创造的独特的信号特征。虽然实验结果较好，但是这种方法依赖于如模数转换器和数字取样示波器等昂贵的设备。此外，他们的分析仅是针对有线设备而言，这种方法对无线网络设备是否可行还存在疑问。

Cristea等人【Fingerprinting Smartphones Remotely via ICMP Timestamps】用线性规划计算ICMP时间戳的时钟偏差，通过观察不同设备的时钟偏差来确定目标。Kohno等人在论文【Remote Physical Device Fingerprinting.被动】中研究根据设备硬件中微小时钟偏移作为识别设备的特征指纹。该技术无需对设备做任何修改，即便目标设备通过多种方式接入互联网，或是测量设备与目标设备相距数千英里，均能达到较好的识别结果。同样基于时钟偏移的识别技术还有【Clock Skew Based Remote Device Fingerprinting Demystified，被动】，与Kohno等人不同的是，【】中作者研究了基于时钟偏差的无线设备识别的局限性，他们利用无线接入点（AP：Access Point）在信标帧中定期发送的时间戳为依据进行识别，消除了测量设备时钟偏移的影响。此外，他们还进行了大量的评估，以探讨不同接入点和测量装置之间时钟偏差的分布及稳定性，发现时钟偏差的波动仅为1ppm。该算法能够消除测量设备的时钟偏斜的影响，使不同设备生成的指纹可比较且具有区分性。

基于硬件的指纹识别技术依赖于某些稳定的特性。【Do You Hear What I Hear，53】中的研究表明网络设备往往具有稳定的时钟偏移，同时也有研究者们通过研究TCP和ICMP的时钟偏移来识别设备【Do You Hear What I Hear，46】。然而，时钟偏移率很大程度上依赖于实验环境【Do You Hear What I Hear，，67】。

与仅使用时钟偏斜作为特征指纹不同的是，Neumann C等人【An empirical study of passive 802.11 Device Fingerprinting】分别以传输速率、帧大小、介质访问时间、传输时间、包内间隔时间等特征作为特征指纹进行802.11设备指纹识别，并比较这些特征在设备识别中的性能差异，为其他的设备指纹识别工作提供了很好的借鉴。Franklin等人研究了在IEEE 802.11兼容设备上运行的固件和设备驱动程序的差异，在设备未被更改的情况下可有效精确地识别无线驱动程序【Passive Data Link Layer 802.11 Wireless Device Driver Fingerprinting】。

Nguyen等人和Li Z等人提出用辐射测量技术为设备认证提供可靠支撑【Device Fingerprinting to Enhance Wireless Security using Nonparametric Bayesian Method】【Securing Wireless Systems via Lower Layer Enforcements】。其原理天线、功率放大器、ADC、DAC等硬件在生产过程中不可能完全相同，因此每个设备具有一系列独一无二的辐射测量量，如振幅、频率、相位等，从中可提取出标识设备的特征指纹。

在论文【AccelPrint: Imperfections of Accelerometers Make Smartphones Trackable】中，作者认为智能手机和平板电脑中的加速度计具有独特的指纹，可以用于设备识别。他们在25个安卓手机和两个平板电脑上验证这一方法，识别精度达到96%。这种方法虽然精度高，但是要求设备上必要要有加速度传感器，且需要某种形式的外部刺激/震动来捕获加速度计的数据，局限性较大。相比而言，本文的工作只需远程捕获网络流量数据，不会被用户发现，操作简单，可实现性高。

Anupam Das【Do You Hear What I Hear? Fingerprinting Smart Devices Through Embedded Acoustic Components】等人认为厂商在制造智能手机的麦克风和扬声器时存在不同程度的缺陷，通过分析不同智能手机上的麦克风和扬声器的声学特征可以对设备进行认证。他们的方法在50个安卓手机上可达到98%的精度。